ABSTRACT

Extensions to latent semantic indexing (LSI), including: phrase processing, creation of generalized entities, elaboration of entities, replacement of idiomatic expressions, and use of data fusion methods to combine the aforementioned extensions in a synergistic fashion. Additionally, novel methods tailored to specific applications of LSI are disclosed.